Grid Related Activities at Caltech

Koen Holtman

Caltech/CMS

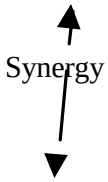
PPDG meeting, Argonne
July 13-14, 2000

1

Executive Summary

Caltech group works on:

- Developing the (Data) Grid
  - Mostly in collaborative efforts
  - Focus on OODBs, regional centres, WAN



- CMS ORCA production
  - To support detector design studies, trigger design studies
  - Real, current needs
  - TBs of data

2

Resources 1

- People
  - People who spend most of their time on grid related work:
  - At Caltech: Julian Bunn, Takako Hickey, Koen Holtman
  - At CERN payed by Caltech: Iosif Legrand, Asad Samar, Mehnaz Hafeez (40%), Philippe Galvez, Gregory Denis
  - Many others spend smaller parts of their time
- Datasets
  - Currently we have
    - ~3 TB of CMS physics data (mostly CMSIM)
    - ~300 GB persistent physics data
- Software
  - CMS software: CMSIM, ORCA, IGUANA
  - Objectivity (C++&Java), Globus
  - [Condor]

3

Resources 2

- Hardware
  - Major CPU:
    - HP X-class (256 CPU) and V-class (128 CPU) systems (Major use for CMS (CMSIM) production planned)
    - Currently using a Condor system at Wisconsin (~50 CPU used during our production)
    - Will build a Linux farm at Caltech
      - - Order in July: 64-128 CPU, Caltech+SDSC
      - - Rising to 256 CPU if Caltech is one of the Tier2 centres
  - Major Storage:
    - HPSS system at CACR (~350 TB)
    - ~1 TB on disk arrays
  - Networking:
    - Many connections -- of particular interest:
      - NTON link to SLAC (2xOC12, 1244 Mbit/s)
      - US- CERN link (45 Mbit/s, probably 155 Mbit/s in Fall)

4

Grid activities 1

- **ORCA production**
  - Production at Caltech
  - Will generate about 1,000,000 events (~1TB) in summer
  - Serves as testbed/motivation for many activities
  - Participation in production elsewhere
    - Involves a lot of file transfers over network
  - Expect to gain valuable experience in running (distributed) production (e.g. To be fed into MONARC)
  - Currently using Condor system at Winsconsin (Vladimir Litvin)
- **MONARC**: see talk by Iosif Legrand

5

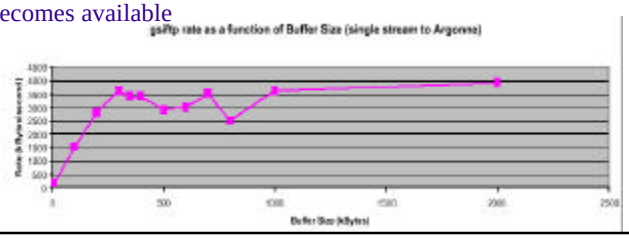
Grid activities 2

- **File based replication** system for distributed ORCA production
  - See talk by Asad Samar
  - Mostly done by CERN-based people (EU DataGrid pilot)
  - First prototype August 2000 (to be used in Fall production)
  - Uses Globus components
- **Object based replication**
  - See (later) talk by Koen Holtman
  - Object (event subset) based replication important for distributed analysis, Tier 2 regional centres and smaller
  - Not a production effort, initial work targeted at demonstration prototype
  - First demo around Autumn 2000
  - Uses Globus components
- Active contacts with Globus team in both efforts

6

Grid activities 3

- **WAN file replication tests**
  - Julian Bunn, Philippe Galvez (network monitoring)
  - Also together with SLAC: see next slide
  - Goals: investigate WAN issues, tune file transport to maximise WAN throughput (TCP window size, routers?)
  - Specific tests (using Globus GSIftp, ORCA files)
    - Caltech-Argonne: tests complete (3.9 Mbytes/s)
    - Caltech-INFN: initial tests complete
  - Will install HPSS enabled version of Globus FTP at CACR as soon as it becomes available



7

Grid activities 4

- **WAN file replication tests 2**
  - For PPDG 100 MB/s milestone, together with Davide Salomoni and Les Cottrell at SLAC
  - NTON link to SLAC (2xOC12, 2x622 Mbit/s = 1244 Mbit/s)
  - Latest results (Julian Bunn):
    - Achieved 380 Mbits/s (47 MB/s) memory-to-memory between Caltech and SLAC over a single OC12 link
    - With a single TCP/IP stream using the iperf tool

8

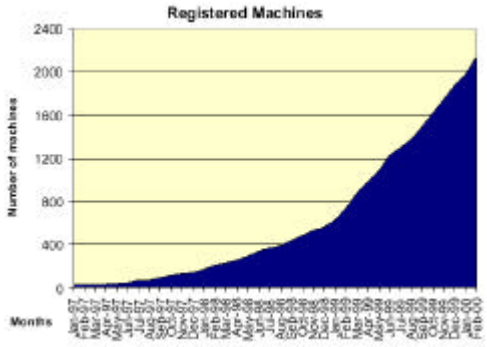
Grid activities 5

- **Distributed resource management**
  - Takako Hickey
  - Research agenda:
    - Fault-tolerant service (special focus on network partition)
    - Data-aware scheduling, Scalable software structure
  - Prototype Execution Service:
    - Submit, monitor, kill jobs individually or as a set
    - Service tolerates all types of failures (network partition, fail-stop, message omissions etc.)
    - Mechanism for placing jobs near data
    - Being tested with ORCA production software (with Vladimir Litvin)
    - Works with 30 CPUs, being debugged for 60 CPUs.
  - Future Work:
    - Study of data-aware scheduling policies using MONARC simulation tool (with Iosif Legrand)
    - Extending prototype for scalability

9

Grid activities 6

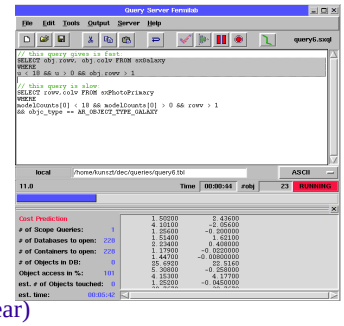
- **VRVS**
  - Philippe Galvez, Gregory Denis at CERN
  - Tools for videoconferencing and collaborative work
  - Part R&D, part production
  - Also performing a lot of WAN network monitoring



10

Grid activities 7

- **ALDAP**
  - Collaboration between Caltech/CMS and JHU/SDSS
    - Both groups are using Objectivity
  - Goal: share experience, technology, develop new ways of organising (object) data
  - SX: data mining tool used in SDSS
    - Use case is similar to end user physics analysis on tag objects and AOD (10 KB) objects
    - Uses an SQL like language
  - Currently installing SX tool at Caltech
  - Will make replica of SDSS data at Caltech (40 GB now, 200 GB in half a year)
  - Doing comparisons between OODBMS and RDBMS (Objectivity and MS SQL Server) for tag/ntuple type analysis



11

Grid activities 8

- **Grid related project coordination**
  - Lot of resources devoted to coordination/synchronisation
    - Between GryPhyn, PPDG, EU DataGrid, MONARC, RD45, ALDAP, ...
  - Harvey Newman, Julian Bunn, Asad Samar

12